DOCUMENT RESUME

ED 270 721 CS 008 458

AUTHOR Atkinson, Michael L.; And Others
TITLE Effects of Training on Accuracy of Decoding Complex Nonverbal Behavior. Working Paper No. 267.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Individualized Schooling.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.r
PUB DATE Aug 79
GRANT OB-NIE-G-78-0217
NOTE 44p.; Report from the Project on Studies of Instructional Programming for the Individual Student.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Audience Analysis; *Behavior Patterns; College Students; *Communication Research; Comparative Analysis; Elementary Education; Elementary School Students; Feedback; Higher Education; *Nonverbal Communication; *Perception; Role Playing; Spontaneous Behavior; Student Behavior; *Training Methods
IDENTIFIERS Accuracy; *Behavior Analysis

ABSTRACT

Two studies investigated the effects of training on how accurately observers (college students) decoded a complex nonverbal stimulus. In the first experiment, observers viewed silent videotapes of 16 third and fourth-grade school children who were listening to an easy or a difficult lesson. Half of the children were responding spontaneously, while the other half had been instructed to pretend that they either understood or did not understand the lesson. Observers were asked to make judgments about the type of behavior they had witnessed (spontaneous or deliberate, and easy or difficult). Some observers were given feedback about their decoding accuracy after each trial, others were given feedback plus an opportunity to discuss the decision among themselves, and a third group served as the control. Results showed that all observers demonstrated a consistent practice effect, but neither feedback nor feedback plus discussion improved decoding accuracy above the level of the control group. The second study examined the effect of practice alone (10 practice trials plus six test trials). Results indicated that when compared to a no-experience control group (who just decoded the six test trials), practice had no effect on decoding accuracy. (Author/HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ED270721

Working Paper No. 267

EFFECTS OF TRAINING ON ACCURACY OF DECODING COMPLEX NONVERBAL BEHAVIOR

by

Michael L. Atkinson, Vernon L. Allen, and Panute Ansusinha

Report from the Project on
Studies of Instructional Programming for the Individual Student

Vernon L. Allen
Faculty Associate

Wisconsin Research and Development Center
for Individualized Schooling
The University of Wisconsin
Madison, Wisconsin

August 1979

CS008458

Published by the Wisconsin Research and Development Center for Individualized Schooling. The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

Center Grant No. OB-NIE-G-78-0217

MISSION STATEMENT

The mission of the Wisconsin Research and Development Center is to improve the quality of education by addressing the full range of issues and problems related to individualized schooling. Teaching, learning, and the problems of individualization are given concurrent attention in the Center's efforts to discover processes and develop strategies and materials for use in the schools. The Center pursues its mission by

- conducting and synthesizing research to clarify the processes of school-age children's learning and development
- conducting and synthesizing research to clarify effective approaches to teaching students basic skills and concepts
- developing and demonstrating improved instructional strategies, processes, and materials for students, teachers, and school administrators
- providing assistance to educators which helps transfer the outcomes of research and development to improved practice in local schools and teacher education institutions

The Wisconsin Research and Development Center is supported with funds from the National Institute of Education and the University of Wisconsin.

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR INDIVIDUALIZED SCHOOLING

Table of Contents

	Page
List of Tables	vii
List of Figures	vii
Abstract	ix
EXPERIMENT 1	
Method	7
Design	7
Stimulus Materials	7
Subjects	8
Apparatus	8
Procedure	9
Method of Analysis	11
Results	13
Conjunctive Accuracy	13
Partial Accuracy	13
Additional Analyses	14
Discussion	19
EXPERIMENT 2	
Method	27
Subjects	27
Apparatus and Stimulus Tape	27
Procedure	27
Results	29
Discussion	33
Reference Notes	35
References	37
Footnotes	39

List of Tables

Table		Page
1	Decoding Accuracy over Trial Blocks as a Function of Initial Accuracy Scores	18
2	Final Accuracy Scores from Jecker et al. (1965)	22
3	Trial-by-Trial Analysis of the Accuracy and Confidence Data	30

List of Figures

Figure		Page
1	Mean accuracy as a function of trial blocks.	15
2	Mean confidence as a function of trial blocks.	25

Abstract

Two studies were conducted to investigate the effects of training on how accurately observers decoded a complex nonverbal stimulus. In the first experiment, observers viewed silent videotapes of school children listening to an easy or a difficult lesson. The stimulus children were either responding spontaneously or deliberately (role playing) and the observers made judgments about the type of behavior they had witnessed (spontaneous or deliberate, and easy or difficult). Some observers were given feedback about their decoding accuracy after each trial; others were given feedback plus an opportunity to discuss the decision among themselves; and a third group served as controls. Results showed that all observers demonstrated a consistent practice effect, but neither feedback nor feedback plus discussion improved decoding accuracy above the level of the control group.

In the second study the effect of practice alone (10 practice trials plus six test trials) was examined. Compared to a no-experience control group (who just decoded the six test trials), practice had no effect on decoding accuracy.

Many studies have investigated the ability of observers to decode nonverbal cues that are associated with various emotional states (e.g., Lanzetta, Cartwright-Smith, & Kleck, 1976; Zuckerman, Hall, DeFrank, & Rosenthal, 1976). In general, the results of these studies indicate that observers are quite accurate in their decoding. Findings are not so unequivocal, however, with more complex nonverbal behaviors. Jecker, Maccoby, Breitrose, and Rose (1964) found that teachers could not accurately discern whether a student was understanding or not understanding a lesson based on the student's nonverbal information alone. Allen and Feldman (Note 1) reported a similar finding in a study using teachers as well as third- and sixth-grade observers. After viewing a videotape of third graders listening to an easy or a difficult lesson, observers were asked to judge how well each student comprehended the lesson. Results indicated that only the children (third and sixth graders) significantly discriminated between easy and difficult lessons.

Contrary to the findings of Allen and Feldman (Note 1), other research (Allen & Atkinson, 1978, Note 2) indicates that adults are able to accurately decode complex nonverbal behavior. College freshmen demonstrated significant discrimination in their ratings of a stimulus person's level of understanding of a lesson; they also were able to estimate the authenticity of the behavior. These results are encouraging, but the level of accuracy in these studies is somewhat disappointing.

Is it possible to improve a person's decoding accuracy in the case

of complex nonverbal stimuli? A procedure resulting in improvement would be particularly valuable in the classroom. Teachers could be trained to increase their skill at decoding nonverbal cues to reasonable levels of comprehension. Difficult sections of a lesson would be detected earlier, and students who require additional instruction could be identified more easily. In general, communication between teacher and student would be greatly enhanced.

In a classic study, Jecker, Maccoby, and Breitrose (1965) reported that they were able to significantly increase decoding ability of teachers. The training procedure involved 6 to 8 hours of rather intensive instruction given to 20 teachers. Various nonverbal cues were suggested as being indicative of different types of behavior. Practice sessions followed the instruction, and the teachers were given feedback about their judgments and encouraged to discuss the interpretation and meaning of the various nonverbal cues they had witnessed. The results indicated that the accuracy levels achieved by teachers were higher after the training sessions.

Although the Jecker et al. (1965) study is often cited as the most convincing demonstration of the effect of training on decoding ability, several problems have been overlooked. The existence of these problems indicates a need for reevaluating the research. The first issue involves a significant pretraining difference between the experimental and control groups. Initially, the decoding accuracy for the control group was significantly higher than that for the experimental group. This fact suggests that any posttraining differences between experimental

and control groups may be due to various idiosyncrasies of the two groups rather than the experimental training session.

The second problem arises from Jecker et al.'s (1965) treatment of this initial difference. They argue that the discrepancy can be controlled by computing improvement scores (final accuracy score minus the initial accuracy score) for each observer. Accordingly, Jecker et al. report that the mean gain in accuracy for the experimental group (7.2%) was significantly different from the mean gain for the control group (-0.5%), and conclude that the training procedure was effective. Unfortunately, this conclusion is misleading. The analysis ignores the fact that the final accuracy score for the experimental group (60.9%) was not very different from the final score for the control group (59.3%). The control group was very stable over the two accuracy measures, and, theoretically, since the control group provides a baseline measure of decoding accuracy, it appears that the "gain" observed in the experimental group merely reflected chance variation in the population parameter. It should also be noted that the final accuracy scores are not much greater than would be expected by chance (50%).

The present research reexamines the problem of improving decoding by training, and presents a more parametric analysis of the problem. The training program devised by Jecker et al. used a combination of instruction, feedback, and discussion. Consequently, it is impossible to determine which of these variables or combination of variables was responsible for the "gain" in accuracy.

The first experiment evaluates the effect of feedback alone and of feedback plus discussion on decoding accuracy. Observers viewed 16

stimulus children who were watching a lesson on a television monitor. This lesson was either easy or difficult for the children, and they were either responding spontaneously (hidden camera) or deliberately (role play). The observer's task was to judge each child's level of comprehension and the authenticity of his or her behavior. In the control condition observers were given no additional information about the children's behavior. Observers in the feedback-alone condition were informed of the correct response after each of the 16 trials. In the feedback plus discussion condition observers also were informed of the correct response; in addition, they were given an opportunity to discuss among themselves their interpretation of the behavior they had witnessed.

The assertion that feedback will improve performance is strongly supported in the literature. Feedback plays a central role in learning theories as a signal concerning the correctness of a response and as a reinforcer in itself. Recently, Powell and O'Neal (1976) demonstrated the importance of feedback in interpersonal perception. Female subjects were asked to give their impressions of a target person either with feedback (subjects were involved in a discussion with the target person and could pose questions to the target) or with no feedback (subjects received the same information as the feedback group but merely watched a discussion). The results indicated that person perception was much more accurate in the feedback condition. Powell and O'Neal (1976) suggest that feedback serves to validate hypotheses about the target that the subject generates during the interpersonal interaction. Subjects who merely view the discussion cannot test their hypotheses and, in the

absence of validated information, it is more difficult to make an accurate assessment of the target person. In a similar manner, feedback should enhance the accuracy with which an observer decodes nonverbal behavior. The observer can use information provided by the feedback to determine which specific nonverbal cues are associated with a particular nonverbal display.

Discussion following feedback is expected to improve decoding accuracy because the discussion is also a source of information. Observers who incorrectly judged a particular segment of behavior have an opportunity to see why they were wrong and will be better prepared to make subsequent judgments. It is difficult to predict a priori if discussion will further enhance decoding accuracy compared to the feedback alone group, but it is reasonable to expect that discussion will not detract from feedback effects. Indeed, Powell and O'Neal's (1976) analysis suggests that more information is available in the feedback plus discussion than in the feedback alone condition. Consequently, the highest decoding accuracy should be observed in the feedback plus discussion group. On the basis of previous research, then, it is expected that both feedback alone and feedback plus discussion will significantly improve decoding accuracy relative to the control condition.

EXPERIMENT 1

Method

Design

The design used in the first study was a 2 (sex of observer) x 5 (training condition) completely randomized factorial. Originally, only three training conditions (control, feedback, and feedback plus discussion) were included in the design, but two additional control groups were added to assess the effect of a possible artifact introduced by the discussion manipulation. The feedback plus discussion group de facto required a longer intertrial interval than the other training conditions, and it is conceivable that any change in performance could be partially attributable to this longer interval. Consequently, a "delayed" control and a "delayed" feedback alone condition were included. The delay simply involved an extension of the intertrial interval to that experienced by the feedback plus discussion group.

Stimulus Materials

Observers viewed silent videotaped recordings of 16 third- and fourth-grade parochial school children (eight males and eight females) who were listening to either an easy or a difficult lesson. Half of the stimulus children (four males and four females) were responding spontaneously; they believed they were in an actual learning situation The other eight children had been instructed to role play or pretend that they either understood or did not understand the lesson. Thus, four categories of behavior were displayed for the observers: (a) spontaneous comprehension, (b) spontaneous noncomprehension, (c) deliberate

comprehension, and (d) deliberate noncomprehension. Each segment of videotape was approximately 2-1/2 minutes in duration, and the 16 segments were presented in random order. Each half of the stimulus tape (eight segments) contained all possible categories of behavior (i.e., one male and one female displaying spontaneous comprehension, one male and one female displaying spontaneous noncomprehension, one male and one female displaying role-played comprehension, and one male and one female displaying role-played noncomprehension). This videotape has been described in detail elsewhere (Allen & Atkinson, 1978).

Subjects

Subjects were 88 students (40 males and 48 females) recruited from an introductory psychology class. They were randomly assigned to one of five training conditions at the time of their arrival: (a) control immediate (7 males and 7 females), (b) control delayed (7 males and 8 females), (c) feedback immediate (8 males and 9 females), (d) feedback delayed (6 males and 8 females), and (e) feedback plus discussion (12 males and 16 females). Observers viewed the videotape in same-sexed groups of three to five persons, depending upon subject availability. At least three observers were present in each of the feedback plus discussion groups.

Apparatus

Observers viewed the videotape on a closed-circuit television system consisting of a Sony 3650 videotape recorder and an RCA Lyceum studio monitor (model #JR345B). An electronic signaling device was used

to record the responses. This signaling apparatus consisted of five booths separated from each other by plywood partitions. A series of nine switches was located at the front of each booth. The first four switches were used to record the observers' decisions concerning the behavior displayed. They were labelled "natural and understanding," "natural and not understanding," "acting and understanding," and "acting and not understanding," respectively. The remaining five switches were inoperative.

Procedure

Observers were conducted to a large viewing room and seated at one of the answering booths. The experimenter explained the content of the stimulus tape. He then instructed the observers to watch each segment of videotape and decide which of the four categories of behavior they believed they had witnessed. The use of the electronic signaling device was explained to the observers and the experimenter continued with specific instructions for each of the training conditions.

For the control conditions, observers were merely told that the experimenter would stop the stimulus tape after each 2-1/2-minute segment of behavior and ask them, via intercom, to indicate their decision by means of the signaling device. After recording their responses, the experimenter would instruct them to reset their switches and the next trial would begin. Observers were asked to watch the stimulus tape and to answer as accurately as possible. For the feedback alone conditions, the experimenter added that he would tell the observers what the actual behavior had been after their responses were recorded.

Subjects were encouraged to use this feedback as a source of information in making subsequent decisions. Again, the experimenter instructed the observers when to reset their switches and prepare for the next trial. For the feedback plus discussion condition, the observers were also informed that the actual behavior would be identified after they had responded. Immediately following feedback the group had approximately 30 seconds to discuss their answers. It was suggested that the discussion might begin by having each group member tell the others what his or her answer had been. Following this, each observer might describe how he or she arrived at that decision. It was hoped that the observers would describe the types of cues they had used in making a decision. The experimenter explained that he would inform them when the discussion period was over and ask them ts reset their switches for the next trial. As before, subjects were encouraged to use the information provided by the feedback and discussion in making subsequent decisions.

In all training conditions the observers were asked to indicate their decisions 10 seconds after the videotape segment had ended. An additional 20 seconds was required to record the responses. At this point, feedback was given to the feedback alone and the feedback plus discussion groups, and the dicussion began for the latter. The control immediate and the feedback immediate groups were instructed to reset their switches and prepare for the next trial. The "delayed" controls and the feedback plus discussion groups were not told to reset their switches until an additional 30 seconds had passed. This procedure produced two intertrial

intervals: 30 seconds for the control immediate and the feedback immediate groups and 60 seconds for the delayed control, delayed feedback, and feedback plus discussion groups.

The experiment was concluded in approximately 1 hour. At the end of the stimulus tape the experimenter returned to the viewing room and informed the observers that they might be asked to participate in a follow-up study and, consequently, would not be debriefed at the present time. After obtaining each observer's phone number, the subjects were excused. Each observer was subsequently contacted by telephone and thoroughly debriefed. (A follow-up study was not conducted.)

Method of Analysis

An observer's response is based on two stimulus dimensions: structure of encoding (spontaneous/deliberate) and content of encoding (comprehension/noncomprehension). If the responses were scored as accurate only when both dimensions are correctly identified, some information would obviously be lost. For example, if the observer chose "acting and understanding" when the stimulus person was, in fact, "acting and not understanding," the observer is partially correct since the structure of encoding was correctly identified. On the other hand, the observer who correctly identified both dimensions has gleaned more information from the videotape than the observer who only identified one dimension correctly.

Consequently, three separate accuracy scores were derived. (a) Accuracy for structure of encoding: an observer's response was defined as accurate if the structure of encoding was correctly identified,

regardless of the response given for content of encoding. (b) Accuracy for content of encoding: an observer's response was considered accurate if the content of encoding was correctly identified, regardless of the response given for structure of encoding. [This particular accuracy score allowed for a direct comparison between the Jecker et al. (1965) research and the present study.] (c) Conjunctive accuracy: a response was defined as accurate only if both dimensions were correctly identified. Note that these three scoring methods are conceptually independent. Scores for the conjunctive accuracy measure are not simply a linear combination of the partial accuracy measures, but represent a different level of decoding accuracy.

Analyses were performed to compare the accuracy obtained on the first eight stimulus persons (block 1) with the final eight stimulus persons (block 2), across conditions. For each type of coding system the maximum score in a given trial block is 8.

Results

Initial analyses were performed to determine if the control immediate and the feedback immediate conditions differed from their respective "delayed" controls. The results indicated no differences on any of the accuracy measures. In view of this finding, the data for the delayed groups were combined with their respective immediate groups.

Conjunctive Accuracy

A 2 x 3 x 2 (sex of observer x training condition x trial blocks) analysis of variance (ANOVA) was performed on the precise accuracy data. Results revealed a main effect for trial blocks, F (1, 82) = 71.77, $p < .001$. Observers were more accurate in their judgments of the final eight stimulus persons (M = 3.96) than they were in the first trial block (M = 2.44). No other effects reached significance. Although a training condition x trial blocks interaction was predicted a priori, a planned contrast was not conducted on the data. Observation of the means for this interaction indicated that the control group demonstrated the largest increase in accuracy from the first to the second trial block.

Partial Accuracy

Similar ANOVA's for accuracy on the structure of encoding and content of encoding measures produced an identical pattern of results. For structure of encoding accuracy, a main effect emerged for trial blocks, F (1, 82) = 146.52, $p < .001$. Observers' accuracy was higher for the second trial block than for the first (M = 5.81 and 3.45, respectively).

Similarly, analysis of the content of encoding accuracy scores produced a main effect for trial blocks, F (1, 82) = 14.97, $p < .001$. Again, observers were more accurate in their final, rather than their initial judgments (M = 5.50 and 4.93, respectively). No other effects reached significance. The pattern of means for the training condition x trial blocks interaction again indicated that the control group manifested the greatest increase in accuracy. Consequently, the planned contrast was ignored.

Additional Analyses

It is possible that trends in the data have been obscured by the use of only two blocks of trials in the analysis. For example, accuracy in the feedback plus discussion group might be a linearly increasing function of trial blocks, but accuracy in the control group might be related to trial blocks by some other function. Although comparing the first eight with the last eight trials produced no reliable between-groups effects, the shapes of the accuracy curves over a larger number of trial blocks may indicate the nature of any learning that occurred. Accordingly, the data were recoded into four trial blocks and a 2 x 3 x 4 ANOVA was conducted on the three types of accuracy scores.

Each of the three analyses revealed only a main effect for trial blocks: conjunctive accuracy, F (3, 246) = 55.01, $p < .001$; structure of encoding accuracy, F (3, 246) = 109.11, $p < .001$; and content of encoding accuracy, F (3, 246) = 52.67, $p < .001$. The results are shown in Figure 1. At the beginning of the experiment, observers were

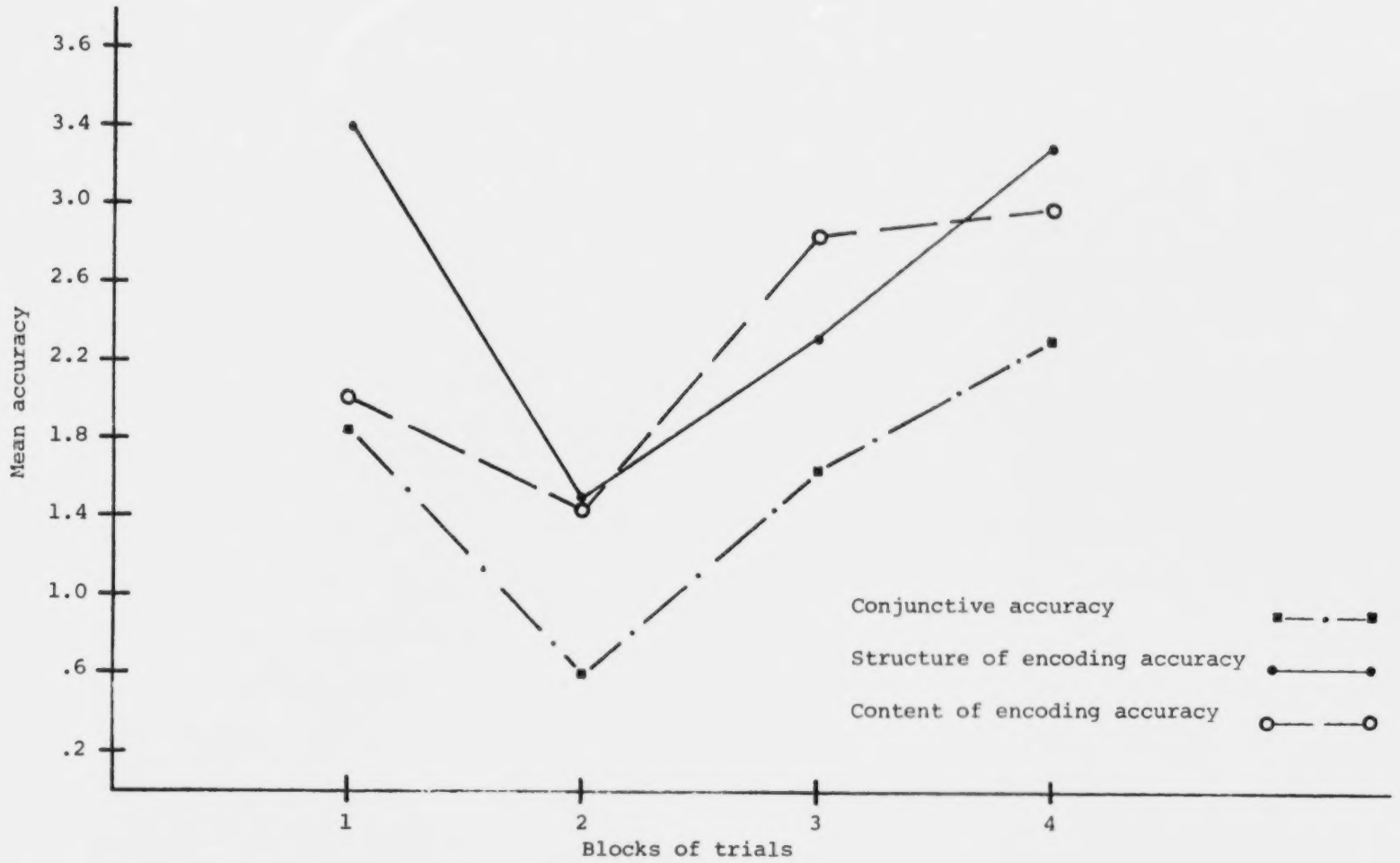

Figure 1. Mean accuracy as a function of trial blocks.

Note. Maximum score for each block = 4.0.

approximately 50% accurate in their judgments (85% with the structure of encoding measure). After observing four more stimulus persons, accuracy was dramatically reduced (e.g., as much as 47% for structure of encoding accuracy). Over the next two trial blocks, accuracy steadily increased for all three measures (an increase of at least 37%). All three curves had significant quadratic and cubic components. In addition, the curves for conjunctive accuracy and content of encoding accuracy demonstrated a significant linear component. No significant differences were found among the training conditions on any of the three accuracy measures.

Finally, we examined the data for possible ceiling effects. Perhaps the task was too easy for the observers and, therefore, training had little effect on decoding accuracy. We designated the observers as high-, medium-, or low-accurate based on their first trial block performance (3 or 4 correct = high; 2 correct = medium; and 0 or 1 correct = low). ANOVA's were conducted on the conjunctive and content of encoding accuracy measures, with initial accuracy (high, medium, or low) as a factor. (We were unable to perform this analysis using the structure of encoding accuracy measure--only 13.6% of the observers scored less than 3 correct on the first trial block and, of these, all scored a 2.) For conjunctive accuracy the ANOVA yielded a main effect for initial accuracy, F (2, 79) = 10.89, $p < .001$; and trial blocks, F (3, 237) = 69.36, $p < .001$. Initial accuracy (high, medium, or low) reflected overall accuracy (M = 1.87, 1.65, and 1.41, respectively), and the effect for trial blocks was identical to the curve shown in Figure 1.

Similar results were obtained with the content accuracy measure.[1] But the most interesting results are shown in Table 1. For both the conjunctive accuracy [F (6, 237) = 16.47, $p < .001$] and the content accuracy [F (6, 237) = 16.69, $p < .001$] scoring methods, an interaction was found between initial accuracy and trial blocks. Observers who were initially high or medium in decoding accuracy showed a decrease from the first to the second trial block, followed by a sharp increase. Low-accurate observers, however, demonstrated very different trends. With the conjunctive accuracy measure, they were relatively stable over the first two trial blocks and then increased in decoding accuracy in much the same manner as the high- and medium-accurate observers. Using the content accuracy measure, the low-accurate observers increased in decoding accuracy quite linearly over trial blocks. Once again, no significant effects were found for the experimental training conditions. The only systematic finding in any of the data analyses was a practice effect: observers simply improved in their decoding accuracy over time.

Table 1

Decoding Accuracy over Trial Blocks as a Function of Initial Accuracy Scores

Initial accuracy	Trial blocks 1	2	3	4
	Conjunctive accuracy			
High	3.18	0.42	1.32	2.54
Medium	2.00	0.58	1.71	2.29
Low	0.84	0.78	1.73	2.29
	Content of encoding accuracy			
High	3.12	1.23	2.52	2.89
Medium	2.00	1.53	2.83	3.17
Low	0.96	1.42	3.11	2.94

Note. Higher numbers represent greater decoding accuracy (maximum score = 4.0).

Discussion

The results indicated that neither feedback nor a combination of feedback plus discussion was sufficient to improve the accuracy of observers' decoding of nonverbal information. If anything, the training conditions reduced decoding accuracy. This result was unexpected, even in view of the weak findings of Jecker et al. (1965).

Let us consider how closely the present study replicates the Jecker et al. research. The major differences between the two studies are found in the training conditions. Observers in the present study were given one of two types of training: feedback alone or feedback plus discussion. This training occurred over a 1-hour period, and observers had no previous experience with the task. On the other hand, observers in the Jecker et al. (1965) study received 6 to 8 hours of intensive instruction, feedback, and discussion.

It may be argued that these training conditions are not comparable across the two studies. The accuracy data indicate, however, that this is not the case. In the Jecker et al. (1965) study observers had initial mean accuracy scores of 56.7% (control and experimental groups combined).[2] The mean final (posttraining) accuracy score was 60.1%--an overall gain of 3.4%. The appropriate comparison in the present study is with the accuracy scores for content of encoding (observers in the Jecker et al. study were asked only to judge whether each student understood or did not understand the lesson). Our observers demonstrated an initial mean decoding accuracy of 50.8% (block 1 of the

analysis using 4 trial blocks). At the end of the 1-hour session the observers were 74.5% accurate in their judgments; an overall gain of 23.7%. (It should be noted that similar results are obtained using the two trial blocks analysis: initial accuracy = 61.6%, final accuracy = 68.8%; a gain of 7.2%.) These data suggest that, at the very least, observers in the two studies were performing at comparable levels before any training, and any improvement in decoding accuracy found in the present study was similar to that reported by Jecker et al. (1965).

A second major difference between the two studies involves the experimental design. Jecker et al. (1965) assessed the effects of training with a pretest-posttest design, whereas we examined training effects as a function of the change in decoding accuracy over trial blocks in the stimulus tape. That is, our procedure involved an ongoing assessment of the training compared to the change-score measure derived by Jecker et al. (1965). Perhaps the effects of training are not evident during the actual training session. It is impossible to dismiss this interpretation completely, but two points argue against it. First, we restate the fact that there was no evidence of a training effect in the Jecker et al. (1965) study, relative to the control group. In this respect, it would seem that the designs used in the present study and in the Jecker et al. (1965) study yielded the same result. Second, we have evidence indicating the decoding accuracy increased simply by exposure to the decoding task; that is, practice improved decoding ability (see Figure 1). Thus, some effects of training (practice) may be observed during the training session.

Jecker et al. (1965) also reported a type of practice effect. In the pretraining test, teachers saw 30 film clips of students responding to questions. The posttraining film contained 90 film clips: 30 repeats (films used in the pretest and the training session); 30 clips of the students from the pretest, but responding to different questions; and 30 novel clips. The results are reproduced in Table 2 (final accuracy scores). Only the main effect for type of film clip was significant ($p < .05$). Regardless of training condition, observers were most accurate with films they had previously seen, moderately accurate with new clips of familiar students, and least accurate with novel films (M = 66.7, 61.4, and 57.4, respectively). Thus, practice (or at least, familiarity[3]) improved observers' decoding ability.

Interestingly, the effects of training in the Jecker et al. (1965) study seem to be stimulus specific (see Table 2): Any benefits of the training experience did not generalize to novel film clips (relative to the control group).[4] This observation suggests a plausible explanation for the lack of training effects in the present study. Our stimulus tape consisted of 16 different stimulus persons. If training effects are stimulus specific, then one would expect little transfer of learning from one trial to the next. Indeed, training may decrease decoding accuracy by providing the observer with inappropriate information about the nonverbal cues indicative of a particular behavior. The data shown in Figure 1 support this contention. Positive effects of practice occurred only after judging the first eight stimulus persons, and accuracy actually decreased between the first and second

Table 2

Final Accuracy Scores from Jecker et al. (1965)

Training condition	Film clip		
	Repeats	New clips, old students	Novel
Experimental	67.8	63.0	56.9
Control	65.6	59.8	57.9

Note. Entries are percentages and reflect posttraining accuracy (not change from initial accuracy).

trial blocks. Table 1 indicates, however, that this decrease only occurred with the initially high- and medium-accurate observers; the low-accurate observers seemed to increase in decoding accuracy in a linear fashion over trial blocks. Presumably, the low-accurate observers have obtained little (if any) information from the first block of stimulus persons and, consequently, did not base subsequent judgments on possibly inaccurate information. Only those individuals who correctly identified the stimulus persons and were informed their judgments were correct have a "validated" [to use Powell & O'Neal's (1976) term] information base to generate hypotheses from during subsequent decoding trials. (This observation accounts for the fact that the control group was often more accurate than either of the training groups.) It should be noted that the linear increase in decoding accuracy found for the low-accurate observers may merely reflect a basement effect--the low-accurate observers can only improve over trials. In this regard, it is interesting to compare the data for the low- and high-accurate observers (Table 1). Although the high-accurate observers were initially three times more accurate than the low-accurate observers, the low-accurate observers demonstrated the fastest increase in decoding accuracy. With one exception the low-accurate observers were consistently more accurate than the high-accurate observers through trial blocks 2, 3, and 4 (the medium-accurate observers fell in the middle). This result is consistent with the suggestion that the high-accurate observers are basing their judgments on inaccurate information.

Data from an earlier study (Allen & Atkinson, Note 1) are also

pertinent to the above discussion. Observers watched silent videotapes and identified both the structure and content of the behavior (similar to the control group in the present study). One group watched the stimulus tape used in the present study. Following the identification of each behavior segment, observers rated their confidence in the judgment they had just given. The data are shown in Figure 2. As can be seen, confidence increased slightly over the course of the stimulus tape. Incidentally, accuracy (by all three scoring methods) followed the same pattern as in the present experiment.[5] These data indicate that although observers show a general decrease in accuracy from the first trial block to the second, they were still confident their answers were correct. Presumably, their confidence is based on information gleaned from the stimulus tape (Attneave, 1959). If the information is incorrect, observers will not be accurate in their decoding of the stimulus tape, but nevertheless may still believe they have correctly identified the behavior. After a number of trials, the correct information is assimilated and accuracy increases. Indeed, a long series of trials (e.g., 90 segments) may be necessary for training procedures (feedback and feedback plus discussion) to have any effect.

Finally, we must be cautious in concluding that trends shown in Figure 1 are due solely to practice. It is quite possible that changes in decoding accuracy over time can be attributed to characteristics of the stimulus persons (and the stimulus order) used in the present study. The data shown in Table 1 and Figure 2 suggest, however, that the "practice" effects are not entirely artifactual. If the time trends merely

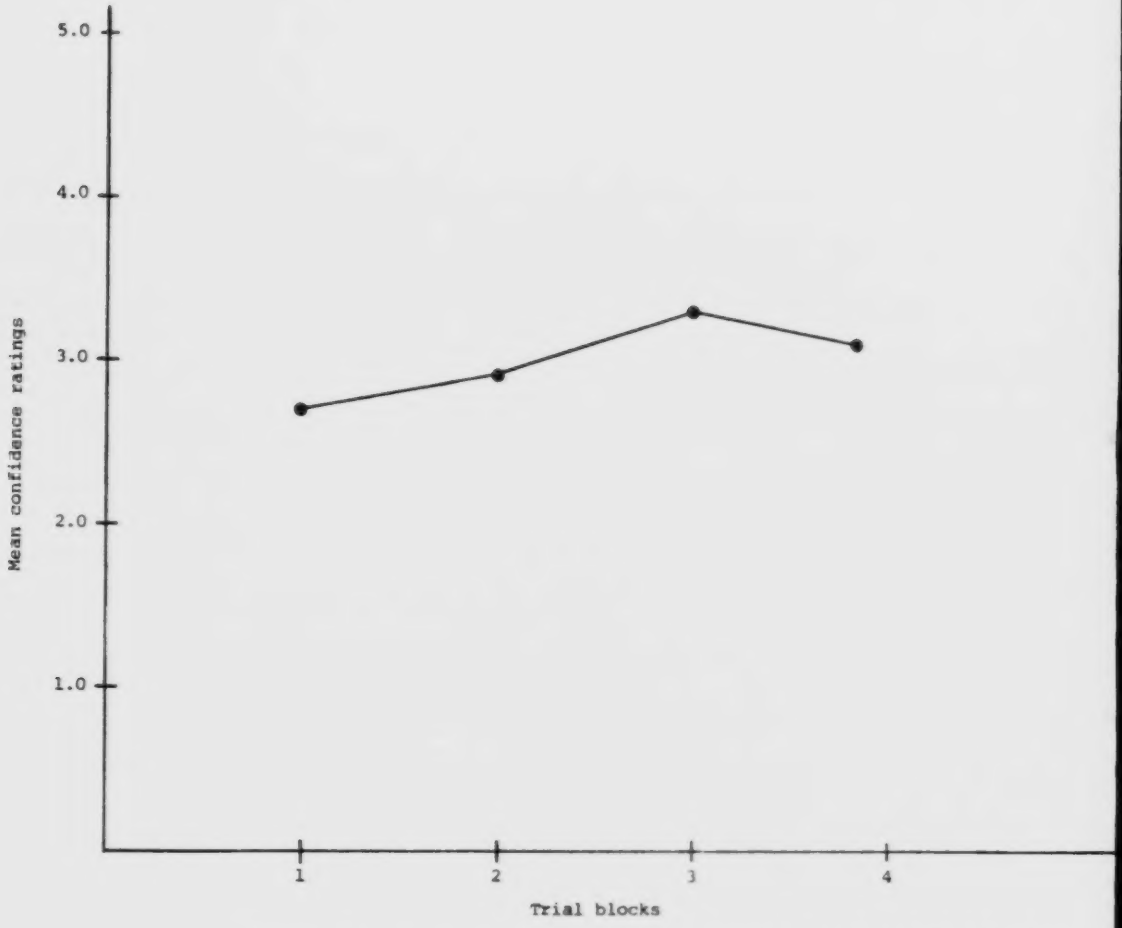

Figure 2. Mean confidence as a function of trial blocks.

reflect task difficulty, for example, then we would not expect the interaction shown in Table 1 (especially that performance of the high-accurate observers would be worse than the low-accurate observers for blocks 2, 3, and 4). In addition, one would expect subjective confidence to be positively related to task difficulty. Figure 2 suggests that the decoding task was easier for the second trial block, not more difficult as the accuracy data (Figure 1) implies.

The results of Experiment 1 indicated that decoding accuracy was not affected by our experimental training procedures. The only consistent finding that emerged from the study was a practice effect--the observers improved over time. The study reported below examines this practice effect in more detail. One group of observers decoded all 16 stimulus persons on the stimulus tape (practice condition), and a second group decoded only the last six (control condition). It was expected that observers with previous decoding experience would be more accurate at decoding the six test segments than observers with no previous experience.

EXPERIMENT 2

Method

Subjects

Subjects were 53 students (20 males and 33 females) recruited from an introductory psychology class. They viewed the videotape in groups of one to five persons (in semi-isolated booths), depending upon subject availability. None had any previous experience with the decoding task.

Apparatus and Stimulus Tape

The apparatus and stimulus tape were identical to that used in Experiment 1.

Procedure

Observers were conducted to a large viewing room and seated at one of the answering booths. The experimenter explained that they would see a series of silent videotapes of children listening to a lesson. The practice group (N = 32) was informed that they would see 16 videotapes, and the control group (N = 21) that they would see six videotapes (the last six segments on the stimulus tape viewed by the practice group). Both groups were asked to judge whether the child in the videotape understood or did not understand the lesson.[6] Observers also were asked to rate confidence for each judgment following the identification of each stimulus tape. The five switches on the right side of the front panel of the signaling device were used for this purpose (1 = not at all confident; 5 = extremely confident). The experimenter then instructed the observers in the use of the electronic signaling device.

When all procedural questions had been answered, the experimenter left the room and started the videotape. Following the decoding session, observers answered a short questionnaire and then were thoroughly debriefed. The experiment was concluded in approximately 1 hour (1/2 hour for the control group).

Results

It was predicted that experienced observers (practice group) should be more accurate in decoding nonverbal behavior than observers with relatively little experience. To test this hypothesis, we compared decoding accuracy on the last six trials for the practice group with the control group's accuracy on these same six segments. Observers were assigned a score ranging from 0 to 6, which indicated how many "test" segments they had correctly identified. Data analysis revealed no difference between the two groups, t (51) = 0.12, n.s. The practice group correctly identified a mean of 4.41 of the six videotape segments (73.5%) compared to 4.38 for the control group (73%). Similarly, no difference was found between the practice (M = 3.36) and control (M = 3.53) groups in reported confidence, t (51) = 1.14, n.s.

We also examined both the accuracy and confidence data on a trial-by-trial basis for the last six trials. These data are shown in Table 3. The practice group held a distinct advantage over the control group on trials one and three, while the control group demonstrated slightly better decoding accuracy on the other four trials. The confidence data, however, did not reflect this pattern. Only on trial two did observers in the practice group report more confidence in their judgments than observers in the control group. Interestingly, the largest difference in reported confidence favoring the control over the practice group was found on trial one--the same trial where the practice group showed the greatest advantage in decoding accuracy over the control

Table 3

Trial-by-Trial Analysis of the Accuracy and Confidence Data

Experimental condition	Trial 1	2	3	4	5	6
3(a) Proportion of observers who correctly identified each stimulus person (last six trials)						
Control	.48	.67	1.00	.71	.86	.62
Practice	.66	.59	.94	.84	.81	.56
3(b) Mean confidence ratings for each of the last six stimulus persons						
Control	3.19	4.00	3.90	3.52	3.67	3.00
Practice	2.69	4.22	3.84	3.47	3.19	2.78

group. The pattern of results shown in Table 3(b) remained essentially unchanged when the confidence data were examined only for those observers who correctly identified the stimulus person on a given trial.

It should be noted that decoding accuracy over trial blocks (practice group) followed the same pattern as in Experiment 1, but the overall accuracy level was higher. Observers in the practice group were 75.25% accurate in the first trial block, dropped to 53% in the second, increased slightly to 56.5% in the third, and attained 78.75% accuracy in the final trial block. We were unable to examine the data by high-, medium-, and low-accurate observers due to the high level of accuracy in the first trial block (most observers correctly identified three or four of the four stimulus persons).[7] Confidence over trial blocks was consistent with the accuracy data (not with the confidence data mentioned in Experiment 1). Reported confidence dropped from the first trial block to the second ($\underline{M}$ = 3.43, 2.91, respectively), and then increased for third and fourth trial blocks ($\underline{M}$ = 3.40, 3.32, respectively).

Finally, the absence of sex of observer as a factor in any current data analysis was suggested by the results of Experiment 1--no significant differences were found between male and female observers. Indeed, an inspection of the means for males and females (both accuracy and confidence data) in the present study revealed almost identical results.

Discussion

The results of Experiment 2 indicated that practice (decoding 10 stimulus persons) did not improve decoding accuracy with respect to a control group who had no previous decoding experience. The confidence data were consistent with this finding--the practice group was no more confident in their judgments than the control group. If anything, observers in the practice group were less confident.

Only one source of data indicates that decoding practice may be of some benefit. Observers in the practice group were approximately 20% more accurate in their judgments of stimulus person number 11 (the first stimulus person seen by the control group) than observers in the control group. One must keep in mind, of course, that any results in the trial-by-trial analysis (Table 3) could be due to any number of factors, e.g., characteristics of a particular stimulus person, the novelty of the task for the control group, etc. Nonetheless, the data presented in Table 3 are interesting. For example, the control group reported much more confidence in their trial one judgment than the practice group. If this initial judgment was difficult for any particular reason (stimulus characteristics, novelty), observers in the control group did not say so. In essence, they were reasonably sure their judgments were correct. The accuracy data indicate that control observers were least accurate on trial one (in fact, at chance level). This is not true for observers in the practice group. Thus, observers

with decoding experience decoded the first stimulus person of the "test" segments more accurately than control observers. This advantage does not appear to be associated with particular characteristics of this initial trial. It is difficult to explain this advantage for the practice group--does experience facilitate decoding ability or merely not inhibit it? The confidence data suggest that observers in the practice group have, at least, acquired some prudence. In general, they were less confident in their judgments than the control group--even when they were more accurate. More research is obviously necessary to determine what effect (if any) practice has on decoding ability.

In summary, the two experiments reported in the present paper do not support the research of Jecker et al. (1965). In Experiment 1 we found that training in the form of feedback and feedback plus discussion did not improve decoding accuracy over the level of a control group that received no training. We did find, however, a consistent practice effect--observers improved in their decoding ability over the course of the stimulus tape. In Experiment 2 we found similar trends over time blocks. More importantly, a control group that saw only the last six children on the stimulus tape performed as well as the experienced observers. Once again, these results are in no way conclusive. Different stimulus materials and orders must also be investigated. The present results do indicate, however, that the potential benefits of decoding training are not as easy to obtain as Jecker et al. (1965) suggest. Training and practice effects are quite subtle, and may not exist for the type of nonverbal behavior we have used (comprehension or noncomprehension for a lesson).

Reference Notes

1. Allen, V. L., & Feldman, R. L. Nonverbal cues to comprehension: Encoding of nonverbal behaviors naturally and by role-play (Working Paper No. 147). Madison: Wisconsin Research and Development Center for Cognitive Learning, 1976.

2. Allen, V. L., & Atkinson, M. L. Identifying spontaneous and deliberate behavior from nonverbal responses (Technical Report No. 419). Madison: Wisconsin Research and Development Center for Cognitive Learning, 1977.

References

Allen, V. L., & Atkinson, M. L. Encoding of nonverbal behavior by high-achieving and low-achieving children. Journal of Educational Psychology, 1978, 70, 298-305.

Attneave, F. Applications of information theory to psychology. New York: Holt, Rinehart, & Winston, 1959.

Jecker, J., Maccoby, N., & Breitrose, H. S. Improving accuracy in interpreting nonverbal cues of comprehension. Psychology in the Schools, 1965, 2, 239-244.

Jecker, J., Maccoby, N., Breitrose, H. S., & Rose, E. Teacher accuracy in assessing cognitive visual feedback. Journal of Applied Psychology, 1964, 48, 393-397.

Lanzetta, J. T., Cartwright-Smith, J., & Kleck, R. E. Effects of nonverbal dissimulation on emotional experience and autonomic arousal. Journal of Personality and Social Psychology, 1976, 33, 354-370.

Powell, R. S., & O'Neal, E. C. Communication feedback and duration as determinants of accuracy, confidence, and differentiation in interpersonal perception. Journal of Personality and Social Psychology, 1976, 34, 746-756.

Zuckerman, M., Hall, J. A., DeFrank, R. S., & Rosenthal, R. Encoding and decoding of spontaneous and posed facial expressions. Journal of Personality and Social Psychology, 1976, 34, 966-977.

Footnotes

[1]Main effect for initial accuracy: F (2, 79) = 7.53, $p < .001$. Main effect for trial blocks: F (3, 237) = 70.76, $p < .001$. Means follow the same pattern as described earlier.

[2]Since no differences were found among the training conditions in the present study, the comparisons reported here present the accuracy data collapsed across this variable. Such a procedure does not misrepresent the Jecker et al. (1965) data. Recall that the initial accuracy scores for their control and experimental groups were 59.8% and 53.7%, respectively. The control observers achieved a final decoding accuracy of 59.3%; the experimental observers, 60.9%.

[3]Jecker et al. (1965) examined familiarity with a second method. Teachers' accuracy scores were compared for their own (home room) versus other students. No differences were found.

[4]Again, we are considering only the final (not the change) accuracy scores. It should be noted that the change score analysis conducted by Jecker et al. (1965) is least appropriate for these novel stimulus tapes (indeed, the analysis should have been conducted only on the repeats). The novel film clips reflect not only the effects of training, but also the fact that both the stimulus children and the stimulus behavior are different.

[5]We were unable to classify the observers as either high, medium, or low in initial accuracy due to the small number of decoders involved (10). However, all decoders on all three accuracy measures correctly identified at least two of the four stimulus persons in the first trial

block. Therefore, the confidence data are applicable to the high- and medium-accurate observers in the present study.

[6]Analysis of the accuracy data for Experiment 1 indicated that the identification of the behavioral structure (spontaneous or deliberate) was relatively easy. Therefore, we simplified the decoding task in Experiment 2 by asking observers to identify only the more difficult behavioral dimension--content of encoding.

[7]We did, however, divide the observers according to their accuracy on the second trial block where a good distribution was available. There was little difference among the high-, medium-, and low-accurate observers with respect to decoding accuracy for the final six stimulus persons. The largest deviation from the overall practice group mean occurred with the medium-accurate observers who correctly identified only 69% of the final six stimulus persons.